39



## SEQUENCE LISTING

#13

```
<110> CHADA, Kiran K.
ASHAR, Hena
TKACHENKO, Alex
ZHOU, Xianjin
```

- <120> HMGI PROTEINS IN CANCER AND OBESITY
- <130> 54615.8001.US02
- <140> US 08/852,666
- <141> 1997-05-07
- <150> US 08/679,529
- <151> 1996-07-12

39

- <160> 54
- <170> PatentIn version 3.1
- <210> 1
- <211>
- <212> DNA
- <213> Artificial Sequence
- <220>
- <223> Rearrangements of 12q15 in human lipomas which disrupt the HMGI-C gene and produce chimeric transcripts
- <220>
- <221> CDS
- <222> (1)..(39)
- <223>
- <400> 1
- <210> 2
- <211> 13
- <212> PRT
- <213> Artificial Sequence
- <220>
- <223> Rearrangements of 12q15 in human lipomas which disrupt the HMGI-C gene and produce chimeric transcripts
- <400> 2
- Pro Arg Gly Arg Pro Arg Lys Trp Pro Gln Gln Val Val 1
- <210> 3
- <211> 39

1

```
<212> DNA
<213> Artificial Sequence
<220>
<223>
      Rearrangements of 12q15 in human lipomas which disrupt the HMGI-C
        gene and produce chimeric transcripts
<220>
<221> CDS
<222>
      (1)..(39)
<223>
<400> 3
cca aga ggc aga cct agg aaa tgg aat act ctg gag cag
                                                                      39
Pro Arg Gly Arg Pro Arg Lys Trp Asn Thr Leu Glu Gln
<210> 4
<211>
      13
<212>
      PRT
<213> Artificial Sequence
<220>
      Rearrangements of 12q15 in human lipomas which disrupt the HMGI-C
        gene and produce chimeric transcripts
<400>
Pro Arg Gly Arg Pro Arg Lys Trp Asn Thr Leu Glu Gln
<210> 5
      39
<211>
<212> DNA
<213> Artificial Sequence
<220>
<223>
      Rearrangements of 12q15 in human lipomas which disrupt the HMGI-C
       gene and produce chimeric transcripts
<220>
<221>
      CDS
<222>
       (1)..(39)
<223>
<400> 5
cca aga ggc aga cct agg aaa tgg ccc ggg ctc cag aag
                                                                      39
Pro Arg Gly Arg Pro Arg Lys Trp Pro Gly Leu Gln Lys
<210>
      6
<211> 13
<212> PRT
<213> Artificial Sequence
```

```
<220>
<223>
      Rearrangements of 12q15 in human lipomas which disrupt the HMGI-C
       gene and produce chimeric transcripts
<400>
Pro Arg Gly Arg Pro Arg Lys Trp Pro Gly Leu Gln Lys
<210> 7
<211> 54
<212> PRT
<213> Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
       encodes transcriptional regulatory domains: chromosome 3
<220>
<221> MISC_FEATURE
<222> (10)..(18)
<223> 9 unspecified amino acids between two sequences
<220>
<221> MISC_FEATURE
<222>
      (36)..(46)
      11 unspecified amino acids between two sequences
<400>
Gln Cys Asn Val Cys Ser Lys Pro Ile Xaa Xaa Xaa Xaa Xaa Xaa
                                                    15
Xaa Xaa Lys Ala Tyr His Pro His Cys Phe Thr Cys Val Met Cys His
           20
                                                30
35
                          40
His Cys Ile Glu Asp Phe
   50
<210>
<211>
      54
<212> PRT
<213> Artificial Sequence
<220>
<223>
      A novel sequence fused to the DNA binding domains of HMGI-C which
       encodes transcriptional regulatory domains:zyxin
```

```
<220>
<221> MISC_FEATURE
<222> (10)..(18)
<223> 9 unspecified amino acids between the two sequences
<220>
<221> MISC_FEATURE
<222>
     (36)..(46)
<223> 11 unspecified amino acids between the two sequences
<400> 8
Lys Cys Ser Val Cys Lys Gln Thr Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                              10
Xaa Xaa Asn Ser Tyr His Pro Gln Cys Phe Thr Cys Val Met Cys His
His Cys Val Asp Asp Tyr
<210>
<211>
     55
<212> PRT
<213> Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
       encodes transcriptional regulatory domains: apteous
<220>
      MISC_FEATURE
<221>
<222>
      (10)..(19)
<223>
      10 unspecified amino acids between the two sequences
<220>
<221> MISC FEATURE
<222>
      (37)..(47)
      11 unspecified amino acids between the two sequences
<223>
<400> 9
Asp Cys Ser Gly Cys Gly Arg Gln Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                                                    15
```

Xaa Xaa Xaa Lys Arq Trp His Ala Ser Cys Leu Lys Cys Tyr Ala Cys

20

25

30

```
35
Ile Tyr Cys Lys Asn Asp Tyr
<210> 10
<211>
     55
<212>
     PRT
<213> Artificial sequence
<220>
     A novel sequence fused to the DNA binding domains of HMGI-C which
<223>
      encodes transcriptional regulatory domains: Lh2
<220>
     MISC_FEATURE
<221>
<222>
      (10)..(19)
      10 unspecified amino acids between the two sequences
<223>
<220>
     MISC_FEATURE
<221>
<222>
      (37)..(47)
<223>
      11 unspecified amino acids between the two sequences
<400> 10
Leu Cys Ala Gly Cys Gly Gly Lys Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Lys Gln Trp His Met Arg Cys Leu Lys Cys Cys Glu Cys
                                            30
Ile Tyr Cys Lys Glu Asp Tyr
<210>
     11
<211>
     53
<212>
      PRT
<213>
     Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
```

encodes transcriptional regulatory domains: Lin-11

```
<220>
<221> MISC_FEATURE
<222> (10)..(19)
<223> 10 unspecified amino acids between the two sequences
<220>
<221> MISC_FEATURE
<222> (37)..(45) ·
<223> 9 unspecified amino acids between the two sequences
<400> 11
Glu Cys Ala Ala Cys Ala Gln Pro Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Xaa Lys Cys Trp His Gln Ser Cys Leu Arg Cys Cys Asp Cys
            20
Arg Ala Pro Met Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Ile Leu
                           40
Cys Lys Thr Asp Phe
    50
<210> 12
<211>
<212> PRT
<213> Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
       encodes transcriptional regulatory domains: RBTN-1
<220>
<221> MISC_FEATURE
<222>
      (10)..(19)
<223> 10 unspecified amino acids between the two sequences
<220>
<221> MISC FEATURE
<222>
      (37)..(48)
<223> 12 unspecified amino acids between the two sequences
<400> 12
Gly Cys Ala Gly Cys Asn Arg Lys Ile Xaa Xaa Xaa Xaa Xaa Xaa
                                   10
```

```
Xaa Xaa Xaa Lys Tyr Trp His Glu Asp Cys Leu Lys Cys Ala Cys Cys
Leu Ile Leu Cys Arg Arg Asp Tyr
<210> 13
<211> 62
<212> PRT
<213> Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
       encodes transcriptional regulatory domains: chromosome 3
<220>
<221> MISC_FEATURE
<222> (10)..(24)
<223> 15 unspecified amino acids between the two sequences
<220>
<221> MISC FEATURE
<222>
      (42)..(54)
<223> 13 unspecified amino acids between the two sequences
<400> 13
Arg Cys Ser Val Cys Lys Glu Pro Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg Phe His Val His Cys Tyr
           20
Arg Cys Glu Asp Cys Gly Gly Leu Leu Xaa Xaa Xaa Xaa Xaa Xaa
       35
                                            45
Xaa Xaa Xaa Xaa Xaa His Ile Leu Cys Lys Thr Cys Asn
   50
                      55
<210>
      14
<211>
     64
<212> PRT
<213> Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
```

## encodes transcriptional regulatory domains: zyxin

```
<220>
<221> MISC FEATURE
      (10)...(25)
<222>
<223>
       16 unspecified amino acids between the two sequences
<220>
<221> MISC_FEATURE
<222>
      (43)..(56)
<223>
       14 unspecified amino acids between the two sequences
<400> 14
Arg Cys Ser Val Cys Ser Glu Pro Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys Asn Phe His Met Lys Cys
Tyr Lys Cys Glu Asp Cys Gly Arg Pro Leu Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Val Leu Cys Met Lys Cys His
<210>
      15
<211>
<212>
      PRT
<213> Artificial Sequence
<220>
      A novel sequence fused to the DNA binding domains of HMGI-C which
<223>
       encodes transcriptional regulatory domains: apterous
<220>
<221> MISC_FEATURE
<222>
      (10)..(20)
<223> 11 unspecified amino acids between the two sequences
<220>
<221>
      MISC FEATURE
<222> (37)..(47)
       11 unspecified amino acids between the two sequences
<223>
<400> 15
Arg Cys Ser Arg Cys Leu Ala Ser Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
```

```
Xaa Xaa Xaa Leu Val Phe His Val Asn Cys Phe Cys Cys Thr Val
          20
                           25
                                            30
40
Ile Tyr Cys Arg Thr His Tyr
   50
                    55
<210> 16
<211> 56
<212> PRT
<213> Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
      encodes transcriptional regulatory domains: Lh2
<220>
<221> MISC_FEATURE
<222> (10)..(20)
<223>
     11 unspecified amino acids between the two sequences
<220>
<221> MISC_FEATURE
<222>
     (38)..(48)
     11 unspecified amino acids between the two sequences
<400> 16
Arg Cys Ala Arg Cys His Leu Gly Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Xaa Leu Val Tyr His Asn Leu Cys Phe Thr Cys Cys Thr
                           25
Leu Val Tyr Cys Arg Leu His Phe
<210>
     17
<211>
     57
<212>
     PRT
<213> Artificial Sequence
<220>
```

٠٠,

```
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
       encodes transcriptional regulatory domains: Lin-11
<220>
<221> MISC FEATURE
<222> (10)..(20)
<223> 11 unspecified amino acids between the two sequences
<220>
<221> MISC_FEATURE
<222> (38)..(49)
      12 unspecified amino acids between the two sequences
<400> 17
Arg Cys Ala Gly Cys Asp Gly Lys Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Xaa Lys Val Phe His Ile Arg Cys Phe Gln Cys Ser Val
           20
                             25
                                                30
35
                          40
Xaa Arg Phe Val Cys Gln Ser Asp Phe
<210> 18
<211>
      56
<212> PRT
<213> Artificial Sequence
<220>
<223> A novel sequence fused to the DNA binding domains of HMGI-C which
       encodes transcriptional regulatory domains: RBTN-1
<220>
<221> MISC_FEATURE
<222>
      (10)..(20)
<223> 11 unspecified amino acids between the two sequences
<220>
<221> MISC FEATURE
<222>
      (38)..(49)
<223>
      12 unspecified amino acids between the two sequences
<400> 18
Asn Cys Ala Ala Cys Ser Lys Leu Ile Xaa Xaa Xaa Xaa Xaa Xaa
                                 10
```

```
Xaa Xaa Xaa Asn Val Tyr His Leu Asp Cys Phe Ala Cys Gln Leu
Xaa Ile Leu Cys Gln Met Asp Tyr
<210> 19
<211> 4
<212> PRT
<213> Artificial Sequence
<220>
<223> consensus residues
<220>
<221> MISC_FEATURE
<222>
      (2)..(3)
<223> 2 unspecified aminio acids between consensus residues
<400> 19
Cys Xaa Xaa Cys
<210> 20
<211> 10
<212> PRT
<213> Artificial Sequence
<220>
<223> consensus sequence of novel sequence fused to the DNA binding
      domains of HMGI-C
<220>
<221> MISC FEATURE
<222> (2)..(9)
<223> The letter "X", or protein "Xaa", indicates any amino acid between
      the consensus residues (which in this case are histidine and
      cysteine)
<400> 20
His Xaa Xaa Cys Xaa Xaa Cys Xaa Xaa Cys
              5
```

```
<210> 21
<211> 4
<212> PRT
<213> Artificial Sequence
<220>
<223> consensus residues
<220>
<221> MISC_FEATURE
<222> (2)..(3)
<223> 2 unspecified aminio acids between consensus residues
<400> 21
Cys Xaa Xaa His
<210> 22
<211> 45
<212> PRT
<213> Artificial Sequence
<220>
<223> Potential transactivation acidic domain encoded by the sequence
       derived from chrosome 15 in ST90-375
<400> 22
Glu Glu Glu Glu His Leu Asn Thr Glu Arg Ser Ser Ala Gly Gly
                                   10
Trp Arg Gly Val Gln Pro Leu Gly Ser Pro Thr Pro Gly Glu Asp His
                               25
Arg Pro Ile Pro Ser Pro Ala Ser Gly Phe Pro Ser Ile
       35
<210> 23
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> Sense primer used to screen for obese mutation
<400> 23
cattctgagt ttgtccaaga tgc
                                                                   - 23
<210> 24
<211> 20
<212> DNA
```

<213>	Artificial Sequence	
<220>		
<223>	Antisense primer used to screen for obese mutation	
<400>	24	
ggtctga	aggc agggagcagc	20
<210>	25	
<211>	23	
<212>		
<213>	Artificial Sequence	
<220> ·		
<223>	Specific oligonucleotide primer synthesized and used to screen	<b>a</b> .
<443>	human total genomic YAC library by the PCT-based method	a
	inditian total genomic the library by the rel-based method	
<400>	25	
	caac aaatgcccac agg	23
agggga	saac aaacgeeeac agg	
<210>	26	
<211>	25	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Specific oligonucleotide primer synthesized and used to screen	а
	human total genomic YAC library by the PCT-based method	
<4.00>	26	
cgtcac	cagg gacagtttca cttgg	25
010	0.5	
<210>	27	
<211>	32 DNA	
<212> <213>	DNA Artificial Seguence	
<213>	Artificial Sequence	
<220>		
<223>	Anchored oligo-dT primer used to synthesize first strand cDNA	
12237		
<400>	27	
	cgac tcactatagt ttttttttt tt	32
3		
<210>	28	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	HMGI-C exon 1 sense primer used in first round of 3' RACE	
<400>	28	٠.
cttcag	ccca gggacaacc	19

<210>	29
<211>	19
<212>	DNA
<213>	Artificial Sequence
<220>	
<223>	antisense adapter primer used in first round of 3' RACE
12207	and the second s
<400>	29
	cgac tcactatag 19
gcaaca	egae ceaetatag
<210>	30
<211>	19
<212>	DNA
<213>	Artificial Sequence
<220>	
<223>	a nested HMGI-C sense primer spanning exon 1 and 2 used to
	reamplify one ml of the PCR reaction
<400>	30
ggaagc	agca gcaagaacc 19
<210>	31
<211>	20
<212>	
<213>	Artificial Sequence
12137	month beganne
<220>	
<223>	primer 375 for performing reverse transcription for the detection
(223)	of chimeric transcripts using novel sequence-specific primers
	of chimeric transcripts using mover sequence-specific primers
-400-	21
<400>	31 toto tgoogcatog 20
CTTCTT	tctc tgccgcatcg 20
<210>	32
<211>	20
<212>	DNA
<213>	Artificial Sequence
<220>	·
<223>	primer 724 for performing reverse transcription for the detection
	of chimeric transcripts using novel sequence-specific primers
<400>	32
	atga taggccttcc 20
J-3~33	
<210>	33
<210>	20
<212>	DNA Autificial Company
<213>	Artificial Sequence
<220>	

<223>	Primer for the novel obtained from lipoma		derived	from	the	chimeric	transcript
<4.00>	33						
cagaag	caga ccagcaaacc					•	20
<210>	34						
<211>			•				
<212>							
<213>	Artificial Sequence						
<220>							
<223>		sequence	${\tt derived}$	from	the	chimeric	transcript
	obtained from lipoma	ST90-375					
<400>	34	:		•			
	tete tgeegeateg						20
-210-	35						
<210> <211>							,
<212>							
<213>	Artificial Sequence					•	
<220>							
<223>	Primer for the novel	sequence	derived	from	the	chimeric	transcript
	obtained from lipoma		•				
4 0 0 .	35						
<400>	agca gtgcaatgtg						20
				•			
	-						
<210> <211>							
<212>							
<213>	Artificial Sequence						
<220>							
<223>	Primer for the novel	sequence	derived	from	the	chimeric	transcript
	obtained from lipoma						
400	,						
<400>	36 atga taggeettee						20
3~3~33							
						•	
<210> <211>	37 23						
<211>	DNA						
	Artificial Sequence						
.000	;					•	
<220> <223>	first strand cDNA sy	nthesis w	ith prime	er 1			
			<b>F</b>	. <del>-</del>			
<400>	37						2.2
atgaat	teet aateeteete tge						23

```
<210> 38
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR amplification with primers 1 and 2
<400> 38
atggatccat gagcgcacgc ggt
                                                                        23
<210>
       39
<211>
      21
<212> DNA
<213> Artificial Sequence
<220>
       21-mer oligonucleotide complementary to the human HMGI-C mRNA
       transcript beginning with the translation initiation codon
<400> 39
gccctcaccg cgtgcgctca t
                                                                       21
<210>
       40
<211>
       20
<212> DNA
<213> Artificial Sequence
<220>
<223>
       20-mer oligonucleotide complementary to the human HMGI-C mRNA
       transcript beginning with the translation initiation codon
<400>
      40
ccctcaccgc gtgcgctcat
                                                                       20
<210>
       41
<211>
       19
<212>
      DNA
<213> Artificial Sequence
<220>
       19-mer oligonucleotide complementary to the human {\tt HMGI-C} mRNA
<223>
       transcript beginning with the translation initiation codon
<400> 41
cctcaccgcg tgcgctcat
                                                                       19
<210>
       42
<211>
       18
<212>
      DNA
<213>
      Artificial Sequence
<220>
<223>
      18-mer oligonucleotide complementary to the human HMGI-C mRNA
```

## transcript beginning with the translation initiation codon

<400> ctcacco	42 gcgt gcgctcat	18
<210><211><212><212><213>	43 17 DNA Artificial Sequence	
<220> <223>	17-mer oligonucleotide complementary to the human HMGI-C mRNA transcript beginning with the translation initiation codon	
<400> tcaccgo	43 egtg egeteat	17
<210><211><212><212><213>	44 16 DNA Artificial Sequence	
<220> <223>	16-mer oligonucleotide complementary to the human HMGI-C mRNA transcript beginning with the translation initiation codon	
<400> caccgc	44 gtgc gctcat	16
<210><211><212><212><213>	45 15 DNA Artificial Sequence	•
<220> <223>	15-mer oligonucleotide complementary to the human HMGI-C mRNA transcript beginning with the translation initiation codon	
<400> accgcgt	45 tgcg ctcat	15
<210><211><212><212><213>	46 21 DNA Artificial Sequence	
<220> <223>	21-mer oligonucleotide complementary to the human HMGI(Y) mRNA transcript beginning with the translation initiation codon	
<400> cttcgag	46 gctc gactcactca t	21

<210>	47	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
(213/	Artificial bequence	
<220>		
<223>	20-mer oligonucleotide complementary to the human HMGI(Y) mRNA	
	transcript beginning with the translation initiation codon	
<400>	47	
	ctcq actcactcat	20
ttegage	acted acted	20
<210>	48	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
12137		
<220>		
<223>	19-mer oligonucleotide complementary to the human HMGI(Y) mRNA	
	transcript beginning with the translation initiation codon	
<400>	48	
	cga ctcactcat	19
oogugo.	3054 0000000	
<210>	49	
<211>	17	
<212>	DNA	
<213>	Artificial Sequence	
	-	
<220>		
<223>	17-mer oligonucleotide complementary to the human HMGI(Y) mRNA	
\2257		
	transcript beginning with the translation initiation codon	
<400>	49	
cgagcto	egac teateat	17
<210>	50	
<211>	17	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	17-mer oligonucleotide complementary to the human HMGI(Y) mRNA	
	transcript beginning with the translation initiation codon	
	eramberips beginning with the translation interaction could	
<400>	50	
gagetee	gact cactcat	17
<210>	51	
<211>	16	
<211>	DNA	
<213>	Artificial Sequence	
<220>		

16-mer oligonucleotide complementary to the human HMGI(Y) mRNA <223> transcript beginning with the translation initiation codon <400> 51 agctcgactc actcat 16 <210> 52 <211> 15 <212> DNA <213> Artificial Sequence <220> <223> 15-mer oligonucleotide complementary to the human HMGI(Y) mRNA transcript beginning with the translation initiation codon <400> 15 gctcgactca ctcat <210> 53 <211> 11 <212> PRT <213> Artificial Sequence <220> consensus sequence for HMGI DNA-binding domains <223> 53 <400> Thr Pro Lys Arg Pro Arg Gly Arg Pro Lys Lys <210> 54 <211> 11 <212> PRT <213> Artificial Sequence <220> <223> consensus sequence for HMGI DNA-binding domains <400> 54 Pro Arg Gly Arg Pro Lys Gly Ser Lys Asn Lys